





Gencore version 5.1.4\_p5\_4578  
 Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 8, 2003, 14:57:17 ; Search time 40 Seconds

(without alignments)  
 2245.273 Million cell updates/sec

Title: US-09-462-517-1

Perfect score: 3490

Sequence: 1 MVQFLGKGQTAGELIHMVTL.....KVSMEVTRPKPLTEAPKA 674

Scoring table: BIOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs., 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
 Maximum Match 100%

Listing first 45 summaries

Database :

1: /SIDS2/gcadata/geneseq/geneseq/geneseq-emb1/AA1980.DAT: \*  
 2: /SIDS2/gcadata/geneseq/geneseq/geneseq-emb1/AA1981.DAT: \*  
 3: /SIDS2/gcadata/geneseq/geneseq/geneseq-emb1/AA1982.DAT: \*  
 4: /SIDS2/gcadata/geneseq/geneseq/geneseq-emb1/AA1983.DAT: \*  
 5: /SIDS2/gcadata/geneseq/geneseq/geneseq-emb1/AA1984.DAT: \*  
 6: /SIDS2/gcadata/geneseq/geneseq/geneseq-emb1/AA1985.DAT: \*  
 7: /SIDS2/gcadata/geneseq/geneseq/geneseq-emb1/AA1986.DAT: \*  
 8: /SIDS2/gcadata/geneseq/geneseq/geneseq-emb1/AA1988.DAT: \*  
 9: /SIDS2/gcadata/geneseq/geneseq/geneseq-emb1/AA1989.DAT: \*  
 10: /SIDS2/gcadata/geneseq/geneseq/geneseq-emb1/AA1990.DAT: \*  
 11: /SIDS2/gcadata/geneseq/geneseq/geneseq-emb1/AA1991.DAT: \*  
 12: /SIDS2/gcadata/geneseq/geneseq/geneseq-emb1/AA1992.DAT: \*  
 13: /SIDS2/gcadata/geneseq/geneseq/geneseq-emb1/AA1993.DAT: \*  
 14: /SIDS2/gcadata/geneseq/geneseq/geneseq-emb1/AA1994.DAT: \*  
 15: /SIDS2/gcadata/geneseq/geneseq/geneseq-emb1/AA1995.DAT: \*  
 16: /SIDS2/gcadata/geneseq/geneseq/geneseq-emb1/AA1996.DAT: \*  
 17: /SIDS2/gcadata/geneseq/geneseq/geneseq-emb1/AA1997.DAT: \*  
 18: /SIDS2/gcadata/geneseq/geneseq/geneseq-emb1/AA1998.DAT: \*  
 19: /SIDS2/gcadata/geneseq/geneseq/geneseq-emb1/AA1999.DAT: \*  
 20: /SIDS2/gcadata/geneseq/geneseq/geneseq-emb1/AA2000.DAT: \*  
 21: /SIDS2/gcadata/geneseq/geneseq/geneseq-emb1/AA2001.DAT: \*  
 22: /SIDS2/gcadata/geneseq/geneseq/geneseq-emb1/AA2002.DAT: \*  
 23: /SIDS2/gcadata/geneseq/geneseq/geneseq-emb1/AA2002.DAT: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length      | DB ID       |
|------------|-------|-------------|-------------|-------------|
| 1          | 3490  | 100.0       | 674         | 20 AAW92953 |
| 2          | 3317  | 95.0        | 674         | 22 ABP59957 |
| 3          | 539   | 15.4        | 111         | 23 AATB7918 |
| 4          | 537   | 15.4        | 207         | 21 AAT53753 |
| 5          | 533.5 | 15.3        | 2070        | 20 AAY04733 |
| 6          | 523.5 | 15.0        | 856         | 21 AAB01383 |
| 7          | 508   | 14.6        | 1373        | 20 AAT04730 |
| 8          | 508   | 14.6        | 2000        | 20 AAY04732 |
| 9          | 505.5 | 100.5       | 20 AAY04731 |             |
| 10         | 491   | 14.5        | 1881        | 20 AAY24205 |

RESULT 1  
 ID AAW92953  
 AAW92953 standard; Protein: 674 AA.  
 AC AAW92953;  
 XX  
 DT 14-MAY-1999 (first entry)  
 XX  
 DE Fly transducisome InaD protein.  
 XX  
 KW InaD; transducisome; fly; insect; signal transduction protein; mutant;  
 KW PDZ domain; modulator; cell surface receptor; ion channel; treatment;  
 KW disorder.  
 XX  
 OS Insecta.

ALIGNMENTS

| Result No. | Query    | Match              | Length | DB ID | Description               |
|------------|----------|--------------------|--------|-------|---------------------------|
| 1          | AAW92953 | standard; Protein: | 674    | AA    | Fly transducisome         |
| 2          | AAW92953 |                    |        |       | Drosophila melanogaster   |
| 3          | AAW92953 |                    |        |       | Human PDZ1 protein        |
| 4          | AAW92953 |                    |        |       | Amino acid sequence       |
| 5          | AAW92953 |                    |        |       | Protein containing        |
| 6          | AAW92953 |                    |        |       | Neuron-associated         |
| 7          | AAW92953 |                    |        |       | Protein containing        |
| 8          | AAW92953 |                    |        |       | Mature Protein containing |
| 9          | AAW92953 |                    |        |       | Amino acid sequence       |
| 10         | AAW92953 |                    |        |       |                           |

XX (AURO-) AURORA BIOSCIENCES CORP.  
 PA (REGC ) UNITV CALIFORNIA.  
 XX  
 PI Mendlein JD, Sierralta J, Sun H, Tsunoda S, Zuker CS;  
 XX DR WPI; 1999-13222/11.  
 DR N-PSDB; AAX02832.  
 XX  
 PT Identifying modulators of signal transduction in cells - used to  
 PT treat signal transduction related disorders  
 PS Claim 9; Page 90-91; 96pp; English.  
 XX  
 CC This invention describes an insect, namely a fly, containing an amino  
 acid mutation in a transducisome protein, InAd, that prevents functional  
 binding of a signal transduction protein. A specific example of the  
 CC mutation is described which occurs in the PDZ domain. The mutated  
 CC protein can be used in a method for identifying modulators of signal  
 transduction, cell surface receptors, ion channels or transducisomes,  
 CC using cells that express the protein or heterologous protein containing  
 a PDZ domain. The chemicals that modulate association of the protein with  
 CC PDZ-binding proteins are useful for treating transducisome-related  
 CC disorders. The transducisome protein can be expressed at high level in  
 E. coli or other hosts, without requiring chemical synthesis or use of  
 CC unnatural amino acid analogues.  
 XX  
 SQ Sequence 674 AA;

Query Match 100.0%; Score 3490; DB 20; Length 674;  
 Best Local Similarity 100.0%; Pred. No. 1. 6e-293; Mismatches 0; Indels 0; Gaps 0;  
 Matches 674; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MVQFLGKQGTAGELIHMVTLDKTGKSFSGCIVREVKSPNPKTGTGIFTKGIVDPSAH 60  
 Db 1 MVQFLGKQGTAGELIHMVTLDKTGKSFSGCIVREVKSPNPKTGTGIFTKGIVDPSAH 60  
 Qy 121 GYMQAKNKFNOEQTTNNASGGQGMQGQGQGMQGAGMNRQSMOKRNFTTASHRQHS 180  
 Db 121 GYGRLKVGDRILSLNGKDVNRNSTEAVIDLIKAEADPKIELEIQTDFDSBQAKSDPRSN 120  
 Qy 181 NYADEDDEDTDRMTGTRTAYEIDRASAGNCNLKQEKDRDKQEDEGYMAKINR 240  
 Db 181 NYADEDDEDTDRMTGTRTAYEIDRASAGNCNLKQEKDRDKQEDEGYMAKINR 240  
 Qy 241 YNMMDKLRLREVORDASKPLGLAGHKDROKMACFVAGYDPPNGALGSVDIKPGEDEVY 300  
 Db 241 YNMMDKLRLREVORDASKPLGLAGHKDROKMACFVAGYDPPNGALGSVDIKPGEDEVY 300  
 Qy 301 NGNVLKNRCLNANASAVFKNVDGDKLUMITSRRKPNPDEGMCKVPIKFPASDETFKIFDQ 360  
 Db 301 NGNVLKNRCLNANASAVFKNVDGDKLUMITSRRKPNPDEGMCKVPIKFPASDETFKIFDQ 360  
 Qy 361 FPKARVQVKERKEGLGIMWYKGKHAEVGSCFISDRESSNAELAGVKYGDMLLVNOV 420  
 Db 361 FPKARVQVKERKEGLGIMWYKGKHAEVGSCFISDRESSNAELAGVKYGDMLLVNOV 420  
 Qy 421 TLESWYDDATGLKRAEGVTTMILTKSEEIAKAKAEEKKKEAKBEEKEEKFQEPATA 480  
 Db 421 TLESWYDDATGLKRAEGVTTMILTKSEEIAKAKAEEKKKEAKBEEKEEKFQEPATA 480  
 Qy 481 EIKPNKKLIELKVKPCKMCHRLLRQKOPCHDWLCLNHPRLSGGOVAARKRLKIFDHICD 540  
 Db 481 EIKPNKKLIELKVKPCKMCHRLLRQKOPCHDWLCLNHPRLSGGOVAARKRLKIFDHICD 540  
 Qy 541 INGTPHVSNTTLVKHOLPHTTVEKAVLTVFADPPELEKFNDFLKKAGKEGLS 600  
 Db 541 INGTPHVSNTTLVKHOLPHTTVEKAVLTVFADPPELEKFNDFLKKAGKEGLS 600  
 Qy 601 PNEIGCTIADIQYQPEIDSKLQRQDITKFGNGALEGIFPQVQYALFGKANGKVSMEV 660  
 Db 601 PNEIGCTIADIQYQPEIDSKLQRQDITKFGNGALEGIFPQVQYALFGKANGKVSMEV 660  
 Qy 661 TRPKPILTEAPKA 674  
 Db 661 TRPKPILTEAPKA 674

RESULT 2  
 ABB5957  
 ID ABB5957 standard; Protein; 674 AA.  
 XX  
 AC ABB5957;  
 XX DT 26-MAR-2002 (first entry)  
 XX DE Drosophila melanogaster polypeptide SEQ ID NO 6663.  
 XX KW Drosophila; developmental biology; cell signalling; insecticide;  
 XX KW pharmaceutical.  
 OS Drosophila melanogaster.  
 XX PN WO200171042-A2.  
 XX PD 27-SEP-2001.  
 XX PF 23-MAR-2001; 2001WO-US09231.  
 XX PR 23-MAR-2000; 2000US19167P.  
 PR 11-JUL-2000; 2000US-0614150.  
 XX PA (PEKE ) PE CORP NY.  
 PI Venter JC, Adams M, Li PWD, Myers EW;  
 XX DR WPI; 2001-656860/75.  
 DR N-PSDB; ABL04060.  
 XX PT New isolated nucleic acid detection reagent for detecting 1000 or more  
 genes from Drosophila and for elucidating cell signalling and cell-cell  
 interactions.  
 PT Disclosure; SEQ ID NO 6663; 21pp + Sequence Listing; English.  
 XX PS The invention relates to an isolated nucleic acid detection reagent  
 CC capable of detecting 1000 or more genes from Drosophila. The invention is  
 useful in developmental biology and in elucidating cell signalling and  
 CC cell-cell interactions in higher eukaryotes for the development of  
 insecticides, therapeutics and pharmaceutical drugs. The invention  
 CC discloses genomic DNA sequences (AB116176-AB130511), expressed DNA  
 CC sequences (ABL0184-ABL16175) and the encoded proteins  
 CC (ABL57737 ABB72072).  
 CC The sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic format directly from WIPO  
 CC at ftp.wipo.int/pub/published\_pct\_sequences.  
 XX  
 SQ Sequence 674 AA;

Query Match 95.0%; Score 3317; DB 22; Length 674;  
 Best Local Similarity 95.3%; Pred. No. 1. 5e-278; Mismatches 7; Indels 24; Gaps 3;  
 Matches 654; Conservative 1; Mismatches 7; Indels 24; Gaps 3;

Qy 1 MVQFLGKQGTAGELIHMVTLDKTGKSFSGCIVREVKSPNPKTGTGIFTKGIVDPSAH 60  
 Db 1 MVQFLGKQGTAGELIHMVTLDKTGKSFSGCIVREVKSPNPKTGTGIFTKGIVDPSAH 60  
 Qy 61 LGGRLKVGDRILSLNGKDVNRNSTEAVIDLIKAEADPKIELEIQTDFDSBQAKSDPRSN 120  
 Db 61 LGGRLKVGDRILSLNGKDVNRNSTEAVIDLIKAEADPKIELEIQTDFDSBQAKSDPRSN 120  
 Qy 121 GYMQAKNKFNOEQTTNNASGGQGMQGQGQGMQGAGMNRQSMOKRNFTTASHRQHS 180

